

PCT09

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

ENTERED

3 <110> APPLICANT: E. I. du Pont de Nemours and Company
 5 <120> TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
 7 <130> FILE REFERENCE: BB-1159
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/744,099
 C--> 10 <141> CURRENT FILING DATE: 2001-06-27
 12 <150> PRIOR APPLICATION NUMBER: 60/093,611
 W--> 13 <151> PRIOR FILING DATE: July 21, 1998
 15 <160> NUMBER OF SEQ ID NOS: 9
 17 <170> SOFTWARE: Microsoft Office 97
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1628
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Zea mays
 24 <400> SEQUENCE: 1
 25 ccgccgcttt ccgagctctc atggcgggcgt ccgcgctctc cctccttgcc gcccccgcac 60
 26 cctcctcctg tggggccatc tccccccagc tccctcgtgg agctccggcc gctgcctccg 120
 27 tcgcttcgcc gtcgctcac tctgctacc tccctccgcg tcccccttcg aggcgccatc 180
 28 ggagtcgctt cgtagccaac gccgctccca ccatgcagcc tccggcggag tccaggggtct 240
 29 ccacggtagt cgatgtcgac ctgcggcacc gtagctatcc gatctacatc ggcgcaggcc 300
 30 tctctgacga gccggacctg ctgcagaggc atgttcatgg taagaggggt ctgggtgggtga 360
 31 ccaacacgac cgtcgcgccg ctttacctgg acaaggtagc atgggcactc acccacaaca 420
 32 acctgaatgt atcagtggaa agcgtgatcc tgcccagcgg tgaaaagtac aaaaatatgg 480
 33 acacgctgat gaaggtgttt gacaaggcag tcgagtcccg ttttgaccgc cgggtgcacat 540
 34 ttgtagcact ggggtgggtgt gtcattgggg acatgtgtgg ttttgacgct gctgcattcc 600
 35 tccggggcgt caatttcata cagataccaa ctactctgat ggcccagggt gattcatctg 660
 36 ttggtgggaa aaccgggatt aaccaccac taggaaagaa cttgattggg gcattctacc 720
 37 agccacaatg tgttctaatt gacacaaata cacttaacac attgcctgac agggagctag 780
 38 cttcaggcat tgccgaggtg gtaaaagtat ggctcataag ggatgcacca ttctttgagt 840
 39 ggcaagagaa gaacatgccg aaattgttag caagagaacc aaatgctttg gcatatgcta 900
 40 tcaagagatc atgtgaaaac aaagctgaag tgggtggcaca agatgagaag gaaagtggcc 960
 41 ttcgagcaac actaaacctg ggtcacacat ttggccatgc tattgagact gggacaggat 1020
 42 atggagcatg gctccatggg gaggtgtctg cagctggaac agttatggca actgacatgt 1080
 43 ctaccgcctt ggggtggata gatgactcca tcagaaaacg tgtggttgac atactaaagc 1140
 44 aagccaaact tcccattgca cctcctgaga ccatgaccgt agagaagttt aaaaacatca 1200
 45 tggctgttga caagaaggtt gctgatggtc tgttgagact catccttctg aaaggaccgc 1260
 46 taggggtgctg tgtatttacg ggggattatg acgggaatgc actcgatgaa accctacatg 1320
 47 cattctgcca caactgagac atcccagttt tggacatcac ttctgtatgc tagtactagg 1380
 48 ctatgtatca atcagagaaa atattgtggt gtacattatc ttactgcct cattggtata 1440
 49 ttttttgtca gaaataatgg tatattgttt taagtgtctg tttgagtttg tagagaacca 1500
 50 tggctaatac actgcggtca gttgttctct agtaataata gtaaagaagc ttactacat 1560
 51 cccttctgtc tgttactgtc aagtaacaat actcgtagag atttaaacca aaaaaaaaaa 1620
 52 aaaaaaaaaa 1628
 54 <210> SEQ ID NO: 2
 55 <211> LENGTH: 438
 56 <212> TYPE: PRT
 57 <213> ORGANISM: Zea mays
 59 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

60 Met Ala Ala Ser Ala Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser
61   1           5           10           15
63 Cys Gly Ala Ile Ser Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala
64           20           25           30
66 Ser Val Ala Ser Pro Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser
67           35           40           45
69 Pro Ser Arg Arg His Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr
70           50           55           60
72 Met Gln Pro Pro Ala Glu Ser Arg Val Ser Thr Val Val Asp Val Asp
73 65           70           75           80
75 Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp
76           85           90           95
78 Glu Pro Asp Leu Leu Gln Arg His Val His Gly Lys Arg Val Leu Val
79           100          105          110
81 Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Thr Trp
82           115          120          125
84 Ala Leu Thr His Asn Asn Leu Asn Val Ser Val Glu Ser Val Ile Leu
85           130          135          140
87 Pro Asp Gly Glu Lys Tyr Lys Asn Met Asp Thr Leu Met Lys Val Phe
88 145          150          155          160
90 Asp Lys Ala Val Glu Ser Arg Phe Asp Arg Arg Cys Thr Phe Val Ala
91           165          170          175
93 Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ala Ala
94           180          185          190
96 Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Leu Met Ala
97           195          200          205
99 Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Pro Leu
100          210          215          220
102 Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln Pro Gln Cys Val Leu Ile
103 225          230          235          240
105 Asp Thr Asn Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly
106          245          250          255
108 Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg Asp Ala Pro Phe Phe
109          260          265          270
111 Glu Trp Gln Glu Lys Asn Met Pro Lys Leu Leu Ala Arg Glu Pro Asn
112          275          280          285
114 Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val
115          290          295          300
117 Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu
118 305          310          315          320
120 Gly His Thr Phe Gly His Ala Ile Glu Thr Gly Thr Gly Tyr Gly Ala
121          325          330          335
123 Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Thr Asp
124          340          345          350
126 Met Ser His Arg Leu Gly Trp Ile Asp Asp Ser Ile Arg Lys Arg Val
127          355          360          365
129 Val Asp Ile Leu Lys Gln Ala Lys Leu Pro Ile Ala Pro Pro Glu Thr
130          370          375          380
132 Met Thr Val Glu Lys Phe Lys Asn Ile Met Ala Val Asp Lys Lys Val

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

133 385          390          395          400
135 Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Cys
136          405          410          415
138 Cys Val Phe Thr Gly Asp Tyr Asp Gly Asn Ala Leu Asp Glu Thr Leu
139          420          425          430
141 His Ala Phe Cys Asp Asn
142          435
144 <210> SEQ ID NO: 3
145 <211> LENGTH: 1358
146 <212> TYPE: DNA
147 <213> ORGANISM: Oryza sativa
149 <400> SEQUENCE: 3
150 gcacgagtag agcgtcgtct tcgtttcgca gctccaccac ctccatggcg gccgcgcct 60
151 cctcctctct gctcgccgcc gcctcgtcgt cctcccgcgc ggcgcccggtg tccgcccgcc 120
152 gcgccccag cgcgagtcgc gccgcggcgg cctccctccc ctgcgcgtct cgcgcctcct 180
153 gcgtcctcc cctccgcgct tccgcagcga ggaccctccg cagccgtgtc gtcgcgagcg 240
154 cggccccgc tatgcagccg ccgcggcggt cgagggtctc cagcgttgtc gacgtcgacc 300
155 tcggcgaccg gagctaccgc atctacatcg gcgcaggcct tctcgacgag cctgacctgc 360
156 tgcagaggca tggtcatggg aagagggttt tgggtgtgac caacaccacc gtgcgcgcgc 420
157 tctacctgga gaagtgacc tgggactca cgcacaacaa ccgaatgtt tctgtggaga 480
158 gcgtgatect gccgcagcgc gagaagtaca aggacatggg cactatgatg aaggttttcg 540
159 acaaggcagt cgagtcccgc ctggaccggc ggtgcacgtt tgttgcgttg ggaggtggcg 600
160 ttattgggga catgtgcggt tttgcagctg ctgcgttctt gcgtggtgtc aatttcatac 660
161 agattcctac tactctgatg gcccgaggtg attcatctgt tggagggaag accggcatta 720
162 accatccatt ggggaagaac ttaattgggg cgttctacca cccacagtgt gtactgatag 780
163 acaccgagac actgaataca ttgcctgaca gggaaactggc ttcaggcata gctgaggtgg 840
164 tgaagtatgg tctcataaga gatgcaccgt tctttgaatg gcaagagaaa aacatgccag 900
165 cattattagc aagagaacca agtgctctgg cctatgctat taagagatcg tgtgaaaaca 960
166 aagctgaagt gggtgctcag gacgagaagg aaagtggctt ccgagcaaca cttaatctcg 1020
167 gccatacatt tggccatgct atagaaacag gaactggcta tggagcatgg ctccatgggg 1080
168 aggctgttgc agctggaaca gttatggcag ctgacatgtc tcaccgcctg gggttgatag 1140
169 acgagtcaat caagaaacgg gcaattgaca tactagagaa agcgaagctt ccaattacac 1200
170 ctccagaggc catgacagtg gagaagttca aaagtattat ggccgttgat aagaaggttg 1260
171 ctgatggatt gctgaggctc atcctcctga aaggacctct gggaagctgt gttttcactg 1320
172 gcgattactg ttcttcgagg tcgacgtgtc gataaatt 1358
174 <210> SEQ ID NO: 4
175 <211> LENGTH: 436
176 <212> TYPE: PRT
177 <213> ORGANISM: Oryza sativa
179 <400> SEQUENCE: 4
180 Met Ala Ala Ala Ala Ser Ser Ser Leu Leu Ala Ala Ala Ser Ser Ser
181 1 5 10 15
183 Ser Arg Ala Ala Ala Val Ser Ala Arg Arg Ala Pro Ser Ala Ser Pro
184 20 25 30
186 Ala Ala Ala Ala Ser Leu Pro Ser Pro Ser Arg Ala Ser Cys Ala Pro
187 35 40 45
189 Pro Leu Arg Ala Ser Ala Ala Arg Thr Leu Arg Ser Arg Val Val Ala
190 50 55 60
192 Ser Ala Ala Pro Ala Met Gln Pro Pro Pro Ala Ser Arg Val Ser Thr

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

193 65          70          75          80
195 Val Val Asp Val Asp Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly
196          85          90          95
198 Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu Gln Arg His Val His Gly
199          100          105          110
201 Lys Arg Val Leu Val Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu
202          115          120          125
204 Glu Lys Val Thr Trp Ala Leu Thr His Asn Asn Pro Asn Val Ser Val
205          130          135          140
207 Glu Ser Val Ile Leu Pro Asp Gly Glu Lys Tyr Lys Asp Met Gly Thr
208 145          150          155          160
210 Leu Met Lys Val Phe Asp Lys Ala Val Glu Ser Arg Leu Asp Arg Arg
211          165          170          175
213 Cys Thr Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly
214          180          185          190
216 Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro
217          195          200          205
219 Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly
220          210          215          220
222 Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr His Pro
223 225          230          235          240
225 Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp Arg
226          245          250          255
228 Glu Leu Ala Ser Gly Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg
229          260          265          270
231 Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Pro Ala Leu Leu
232          275          280          285
234 Ala Arg Glu Pro Ser Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu
235          290          295          300
237 Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg
238 305          310          315          320
240 Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr Gly
241          325          330          335
243 Thr Gly Tyr Gly Ala Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr
244          340          345          350
246 Val Met Ala Ala Asp Met Ser His Arg Leu Gly Trp Ile Asp Glu Ser
247          355          360          365
249 Ile Lys Lys Arg Ala Ile Asp Ile Leu Glu Lys Ala Lys Leu Pro Ile
250          370          375          380
252 Thr Pro Pro Glu Ala Met Thr Val Glu Lys Phe Lys Ser Ile Met Ala
253 385          390          395          400
255 Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys
256          405          410          415
258 Gly Pro Leu Gly Ser Cys Val Phe Thr Gly Asp Tyr Cys Ser Ser Arg
259          420          425          430
261 Ser Thr Cys Arg
262          435
264 <210> SEQ ID NO: 5
265 <211> LENGTH: 1643

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

266 <212> TYPE: DNA

267 <213> ORGANISM: Glycine max

269 <400> SEQUENCE: 5

```

270 gcacgagccc ggagtgggtt ttgtgttcaa ggaagacagg cgcaggcaca ggcacatcat    60
271 gttattgtta ttgttggctt cttcactcag caaccaccaa accacaccgt cgcacccatt   120
272 ccatcgctga ttcccccaat ggcttccact gccaccaatt tctctctttc tctctgcgcc   180
273 aaccaacaaa ctccaatccc caaacctctt ttcttctcca ataacaacca ttgtcacttc   240
274 aactctaata ataattgggc ctgggcctct gtttccacct ctgcgaagtc aaggatatgc   300
275 gccacctcct ctcaagttaa ggatccctcc gcagcaaaat ccgaaccgcg tcttcccacc   360
276 atcgtcgaag tcgatttggg tagccggagc tatcctattt acatcggatc cgggttacta   420
277 aaccaaccgc actatctcca gaggcattgt catggaaaga gagtccatgt tgtaactaac   480
278 gaaaccgttg cgccacttta tctagacaag gttgttgatg ctttgacaag gggaaaccgc   540
279 aatgtttctg tggagagtgt aattttacct gatggtgagc agtacaagga catggatact   600
280 cttatgaaag tctttgacaa ggccatcgag tcgcggctcg accggcggtg tacatttgtt   660
281 gctcttgggt gtggtgtgat tggcgacatg tgtggctttg ctgcctctgc cttcctacgt   720
282 ggtgttaatt ttattcagat tcctacaact gtgatggcac aggtcgattc ttcagttggt   780
283 ggaaaaactg ggataaatca ccgccttggg aagaacatga tcggtacctt ttaccaacct   840
284 cagtgtgtgc ttatagacac agacacatta aatacgctac cggataggga actggcatca   900
285 gggctagcag aggttataaa gtatgggctc attagggatg cagagttttt tgagtggcaa   960
286 gagaaaaata tgcacttatt attggcaaga gatcctagtg taatggcata tgctataaag  1020
287 cgatcttgtg aaaacaaggc tgaggttgtg tccttagatc agaaggaaaag tggactgagg  1080
288 gcaacattga acttgggtca tacatttggg catgcaatag aaactggggt aggctatggg  1140
289 cagtggcttc atggagaggc tgttgacagt ggcacggtaa tggctgttga catgtcatat  1200
290 cgcctaggtt ggattgatga ttctcttgtg aaacgagttg gagacatttt aaaacaggct  1260
291 aagttacceca cagccctccc tgagaccgtg actgtggaca tgtttaaatc tgtcatggca  1320
292 gtggataaga aggtagcaga tgggttgcta aggccttatcc ttctaaaggg tcctctaggg  1380
293 aattgtgttt tcacagggga ttatgacaga aaggctctag acaatacgct tcgtgcattc  1440
294 tgtaaatcct gatcttcttg ctccattgtt ctattgcagc tttatccaat tccttttagt  1500
295 gagtttttgt atgttgttat ataacctcac ttgtattttt cttatgtact ttggatggca  1560
296 accatggttc ccattatata cgcaagagta tattatttga aattgaggca tgatcaataa  1620
297 ttattaaaaa aaaaaaaaaa aaa                                1643

```

299 <210> SEQ ID NO: 6

300 <211> LENGTH: 437

301 <212> TYPE: PRT

302 <213> ORGANISM: Glycine max

304 <400> SEQUENCE: 6

```

305 Met Ala Ser Thr Ala Thr Asn Phe Ser Leu Ser Leu Cys Ala Asn Gln
306   1             5             10             15
308 Gln Thr Pro Ile Pro Lys Pro Ser Phe Phe Ser Asn Asn Asn His Leu
309             20             25             30
311 His Phe Asn Ser Asn Asn Asn Trp Ala Trp Ala Ser Val Ser Thr Ser
312             35             40             45
314 Arg Lys Ser Arg Ile Cys Ala Thr Ser Ser Gln Val Met Asp Pro Ser
315             50             55             60
317 Ala Ala Lys Ser Glu Pro Ala Leu Pro Thr Ile Val Glu Val Asp Leu
318             65             70             75             80
320 Gly Ser Arg Ser Tyr Pro Ile Tyr Ile Gly Ser Gly Leu Leu Asn Gln
321             85             90             95
323 Pro Asp Tyr Leu Gln Arg His Val His Gly Lys Arg Val Leu Val Val

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/744,099

DATE: 08/01/2001

TIME: 18:20:00

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:13 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD